

# Les données du SAMU comme moyen de surveillance de la santé de la population

Naprous, Alexandre, Université de Bordeaux & BPH INSERM U1219, Bordeaux (Orateur)

Avalos-Fernandez, Marta, INRIA SISTM & Université de Bordeaux & BPH INSERM U1219, Bordeaux

Pradeau, Catherine, CRRA 15 - SAMU 33 - SMUR de Bordeaux CHU de Bordeaux & Université de Bordeaux & BPH INSERM U1219, Bordeaux

Lagarde, Emmanuel, Université de Bordeaux & BPH INSERM U1219, Bordeaux

Gil-Jardiné, Cédric, Service des Urgences du CHU de Bordeaux & Université de Bordeaux & BPH INSERM U1219, Bordeaux

prénom.nom@u-bordeaux.fr

Thématique : Analyse des données

Résumé :

*Les données enregistrées lors d'appels aux SAMU en Gironde constituent une source d'information riche (données temporelles, spatiales, cliniques, textuelles). L'application de méthodes d'apprentissage aux rapports disponibles en textes libres permet de classer ces appels en fonction des pathologies ou des symptômes en cause. Nous avons également développé un outil de visualisation des tableaux de bord afin de réaliser une analyse exploratoire dynamique et spatiale de ces données.*

Mots clés : datamining, réseaux de neurones transformers, analyse exploratoire de données volumineuses, données de Gironde, analyse spatiale, dashboard

## 1. Introduction

Les centres de régulation médico-sanitaire des urgences (SAMU) ont pour mission de recevoir et traiter les appels d'urgences médicales. Le SAMU assure une écoute médicale permanente ; détermine et déclenche la réponse la plus adaptée à la nature des appels ; s'assure de la disponibilité des moyens d'hospitalisation adaptés à l'état du patient ; organise, le cas échéant, le transport dans un établissement en faisant appel à un service public ou à une entreprise privée de transports sanitaires ; assure le suivi du patient, le suivi des décisions et des effecteurs engagés par la régulation médicale ; et puis veille à l'admission du patient.

Avec un million d'appels par an, le SAMU de la Gironde est le 2ème centre d'appels d'urgence de France. Son système d'information permet l'archivage exhaustif de l'ensemble des appels et de leur traitement depuis 2005. Cette base de données est de grande dimension, par le nombre d'enregistrements archivés, mais aussi par l'étendue des informations qu'elle contient. En effet, pour chaque appel sont renseignés la provenance de l'appel, le profil de l'appelant, le type du lieu d'intervention, les circonstances, le motif de recours au SAMU, le devenir du patient (domicile, transport, décès) et les actions déclenchées par l'appel (conseil, nature des moyens mis en œuvre). Une autre catégorie de renseignements est donnée sous forme de textes libres. On y trouve enfin des informations permettant de géolocaliser la victime, ainsi que les paramètres vitaux mesurés par les

appelants et les intervenants. Ces données sont une source d'information qui peut être utilisée pour mettre en place une surveillance épidémiologique en temps réel.

Notre projet a consisté à extraire des informations structurées de ces données en texte libre et à mettre en place des outils permettant une analyse exploratoire temporelle et spatiale d'indicateurs de santé de la population de la Gironde. Les objectifs étaient d'observer l'évolution de cette répartition des symptômes évoqués dans la période pré, per et post épidémique, à la recherche d'un pattern spécifique qui signerait de manière précoce le début de l'épidémie et de construire un outil de visualisation des données de santé issue du CRRA15-SAMU de la Gironde.

## **2. Méthodologie**

### **1. Extraction de données issues des rapports en texte libre**

Dans une première partie, nous avons cherché à extraire de nouvelles informations à partir des données des dossiers de régulation médicale notées en texte libre par les assistants de régulation médicale et les médecins du centre de régulation. On en compte en moyenne 300.000 par an depuis 2009, soit un total de plus de trois millions entre 2009 et 2021.

Nous avons utilisé des méthodes de classification supervisée par réseaux de neurones de type Transformers pour classifier ces rapports en fonction de la raison de l'appel ou de la pathologie (état grippal, les douleurs thoraciques, avc, violence, etc...)

A l'aide d'une recherche par mots clés, nous avons également extrait les délais nécessaires à l'arrivée des secours sur les lieux et les tensions artérielles transmises lors de l'appel par les professionnels de santé ou secouristes sur place.

### **2. Données géographiques**

Pour chaque dossier sont renseignés la commune de provenance de l'appel et le Code Officiel Géographique, ainsi que l'adresse de localisation du patient transcrite en texte libre par le soignant. Pour tirer parti au mieux des données géographiques, nous avons déterminé les coordonnées GPS des adresses en utilisant l'API disponible sur « [adresse.data.gouv.fr](https://adresse.data.gouv.fr) » dans un script R.

### **3. Outils d'analyses pour valoriser ces données**

Finalement, nous avons agrégé les données pour mener des analyses exploratoires concernant la santé de la population de la Gironde. Pour cela, nous avons développé un outil de visualisation (*dashboard*) permettant l'exploration géographique de ces données.

Ce dashboard comporte deux éléments principaux : une carte interactive de la Gironde où sont affichées les données et des panneaux de sélection de variables à visualiser, des périodes temporelles à filtrer et du mode d'affichage des données. Les données sont traitées en temps réel en fonction des différents choix utilisateur, offrant une large gamme de possibilité de visualisation et d'exploration. L'outil a été développé sous R et utilisent les bibliothèques `Rshiny` pour le dashboard et `leaflet` pour la cartographie.

Cet outil a permis l'analyse des provenances d'appels pour différents symptômes évocateurs de Covid19 et la comparaison avec d'autres périodes épidémiques passées.

Dans le cadre de l'analyse portant sur l'épidémie de Covid19, nous avons aussi effectué une analyse statistique de la répartition spatiale des lieux de provenance des appels au 15. A

cette fin, nous avons utilisé l'indice d'autocorrélation spatiale (I de Moran) qui mesure si l'état d'une commune au regard du nombre d'appels pour un symptôme donné est plus ou moins corrélé avec l'état des communes plus ou moins proches. Cette analyse n'a pas permis de mettre en évidence de signaux spécifiques, suggérant l'inadaptation de ces données à l'alerte et au suivi d'une épidémie.

L'outil de visualisation spatiale a également été appliqué à l'analyse de la rapidité de la prise en charge des patients sur les lieux avec des cartes de temps de trajet moyen par commune des secours. Il a aussi permis l'étude de la répartition géographique d'indicateurs de santé de la population comme la tension artérielle, suggérant un lien avec les profils socio-économiques des territoires concernés.

### **3. Originalité / perspective**

L'originalité du projet reposait dans l'extraction d'une nouvelle source de données de santé publique disposant d'informations fines à la fois géographiques et spatiales. Par un petit nombre de cas d'usage, il a permis d'initier une réflexion sur leur utilisation comme évaluateur de la santé des populations et l'utilisation que pourrait en avoir les autorités sanitaires.

Ce travail a également permis de créer un outil de visualisation inédit qui permet une analyse géographique fine et qui pourra encore être adapté à de nombreuses autres problématiques de santé des populations.

### **Références**

C Gil-Jardiné, G Chenais, C Pradeau, E Tentillier, P Revel, X Combes, M Galinski, E Tellier and E Lagarde. Trends in reasons for emergency calls during the COVID-19 crisis in the department of Gironde, France using artificial neural network for natural language classification. Scand J Trauma Resusc Emerg Med, 2021 Mar 31;29(1):55.